Journal of Experimental Psychology: Learning, Memory, and Cognition

The Development of Automaticity in Short-Term Memory Search: Item-Response Learning and Category Learning

Rui Cao, Robert M. Nosofsky, and Richard M. Shiffrin

Online First Publication, December 8, 2016. http://dx.doi.org/10.1037/xlm0000355

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/16/$12.00 http://dx.doi.org/10.1037/xlm0000355

The Development of Automaticity in Short-Term Memory Search: Item-Response Learning and Category Learning

Rui Cao, Robert M. Nosofsky, and Richard M. Shiffrin
Indiana University Bloomington

In short-term-memory (STM)-search tasks, observers judge whether a test probe was present in a short list of study items. Here we investigated the long-term learning mechanisms that lead to the highly efficient STM-search performance observed under conditions of consistent-mapping (CM) training, in which targets and foils never switch roles across trials. In item-response learning, subjects learn long-term mappings between individual items and target versus foil responses. In category learning, subjects learn high-level codes corresponding to separate sets of items and learn to attach old versus new responses to these category codes. To distinguish between these 2 forms of learning, we tested subjects in categorized varied mapping (CV) conditions: There were 2 distinct categories of items, but the assignment of categories to target versus foil responses varied across trials. In cases involving arbitrary categories, CV performance closely resembled standard varied-mapping performance without categories and departed dramatically from CM performance, supporting the item-response-learning hypothesis. In cases involving prelearned categories, CV performance resembled CM performance, as long as there was sufficient practice or steps taken to reduce trial-to-trial category-switching costs. This pattern of results supports the category-coding hypothesis for sufficiently well-learned categories. Thus, item-response learning occurs rapidly and is used early in CM training; category learning is much slower but is eventually adopted and is used to increase the efficiency of search beyond that available from item-response learning.

Keywords: memory search, recognition, response time, short-term memory, long-term memory

Memory researchers tend to focus on either episodic memory for recent events or long-term learning, but of course learning can and does occur even in episodic short-term-memory (STM) tasks. In this article, we study learning that occurs in certain variants of the memory search task popularized by Sternberg (1966): A short list of items is presented in sequence, followed by a test probe that is either a "target" from the list or a "foil" not from the list. Performance is generally highly accurate, and the primary data of interest are the way response time varies with the size of the list, termed the set-size effect. In the typical paradigm, the stimuli that are targets on one trial are foils on another, and vice versa, a method termed *varied mapping* (VM) by Schneider and Shiffrin (1977). In such a paradigm, response time rises more or less linearly as set size increases (errors are low but also increase with set size). Sternberg (1966, 2016) proposed serial-exhaustive search of the list set to explain the findings observed under his conditions of testing. Other research that increased presentation rates and shortened the delay between list presentation and test as provided evidence for alternative models, in which response time is largely determined by lag or recency of the test item (e.g., McElree & Dosher, 1989; Monsell, 1978; Nosofsky, Little, Donkin, & Fific, 2011).

This article uses a memory search paradigm but focuses on a different question: What sorts of long-term learning contribute to performance in different variants of the task? In the typical VM paradigm, an item might serve as a target on one trial, thereby producing some learning that that item is a target. But on other trials the same item might serve as a foil, pushing learning in the other direction. This inconsistency would lead to memory interference (e.g., Nosofsky, Cox, Cao, & Shiffrin, 2014) and/or suppress long-term learning. As a result, in a VM paradigm, performance tends to rely primarily on retrieval of items from only the current list held in short-term memory, and the set-size effect tends to retain its magnitude as training continues.

However, VM is not the only memory search paradigm. Shiffrin and Schneider (1977), in their hybrid memory search/visual search procedure, also used *consistent mapping* (CM). In CM the targets remain targets on every trial, and foils remain foils on every trial. As training proceeds, it is evident that some form of learning has taken place: Subjects perform faster with fewer errors and, most important, the slope of the set-size function drops toward zero, a result suggesting an increasing reliance on a process other than retrieval from a list held in STM (see also Logan & Stadler, 1991). Indeed, CM performance is often considered a hallmark example of the development of forms of automaticity in cognition (for extensive converging evidence, see Schneider & Shiffrin, 1977, and Shiffrin & Schneider, 1977). Thus, it is important to develop a deeper understanding of the mechanistic bases for CM learning.

In a CM paradigm, at least two types of long-term learning can take place. First, the observer may learn that a given item always has the

Rui Cao, Robert M. Nosofsky, and Richard M. Shiffrin, Department of Psychological and Brain Sciences, Indiana University Bloomington.

This work was supported by Air Force Office of Scientific Research Grant FA9550-14-1–0307 to Robert M. Nosofsky. Our thanks to Gordon Logan for helpful suggestions on an earlier version of this article.

Correspondence concerning this article should be addressed to Robert M. Nosofsky or Rui Cao, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405. E-mail: nosofsky@indiana.edu or caorui@indiana.edu

response “target” (or always has the response “foil”). We refer to this form of learning as “item-response learning,” a mechanism that serves as a core component of Logan’s (1988) influential theory of the development of automaticity. Nosofsky, Cao, Cox, and Shiffrin (2014) examined cases in CM and VM training in which the same foil was repeated across consecutive trials. Whereas this manipulation led to massive interference in VM, if anything there was slight facilitation in CM. This facilitation in CM is consistent with the view that observers were forming memories of the actual item-response mappings (i.e., target vs. foil), rather than responding solely based on a simple increase in stimulus familiarity (for converging evidence in the domain of hybrid visual/memory search, see Wolfe, Boettcher, Josephs, Cunningham, & Drew, 2015).

A second potential form of learning in CM is learning that all the target items are the members of a single category with a common (perhaps implicit) label and that all the foil items are members of a separate category (e.g., Cheng, 1985; Logan & Stadler, 1991; Schneider & Shiffrin, 1985; Shiffrin & Schneider, 1977). The latter kind of learning is typical of concept and category learning that takes place in life after a great deal of experience—a category label encodes a category, and the encoding involves more than standing for a list of the category members. Examples of such categories include animal, letter, and number. In the CM paradigm, once observers learn the categories, they can then respond “old” or “new” by determining the category to which the test item belongs, without the need to search the memory set presented on the current trial. Use of learned categorization could thereby eliminate the memory search set-size effect.

Crucially, in their hybrid visual memory search paradigms, Shiffrin and Schneider (1977) observed greatly enhanced CM performance (compared to VM performance) even in cases in which targets and foils were grouped into arbitrary categories, albeit after a great deal of training. Likewise, in pure memory-search paradigms, Nosofsky, Cox, et al. (2014) also used arbitrary categories and observed similar forms of enhanced CM performance, even at early stages of practice. In the present article, we examine memory search both for categories made up initially of arbitrarily related items and for existing well-learned categories.

In sum, in a CM paradigm, performance could come to rely on either item-response learning or category learning. In the former, given some test item, a response could be based on the learned response (“old” or “new”) to that particular item. In the latter, the item could be coded as a member of a newly learned category, and the response based on a response to the category label. Note that in both cases, the speed of the response would not depend on the size of the current memory set.

In the present research, our key theme is to try to distinguish between the two types of learning mechanisms by making use of *categorized varied mapping* (CV). In this paradigm, there are two groups of items; all items within each group are always assigned to the same role on a given trial, but the roles switch from trial to trial. For example, for ease of description, suppose the two sets are letters and numbers: On one trial the memory set might be chosen from the numbers and a foil from the letters, and on another trial the memory set from the letters and a foil from the numbers. Thus, the memory set and foil are always from different categories, but the categories switch roles across trials. Obviously, item-response learning cannot take place in CV, because the mapping of a given item to a response changes from trial to trial. The critical point is that category learning can take place in CV because the items of a given category are consistently assigned together, whether the response is a target or foil. Thus, if the test item belongs to the same category as the memory-set items, then it must be a target, whereas if it belongs to the alternative category, then it must be a foil.

If category encoding is not learned, then because item associations cannot be used, the CV results should align closely with VM, with both showing large set-size effects. But to the extent that category encoding is learned and used, then set-size effects should drop. To repeat the logic, the observer need only note the category of the memory set and then check to see whether the category of the test item matches. This strategy would produce a binary decision unrelated to the number of items presented.

Thus, in the present research we tested between the item-response-learning and category-learning possibilities by conducting memory search tasks involving VM, CM, and CV training. Although Shiffrin and Schneider (1977) included a CV condition in one of their studies, it was conducted in the context of a hybrid memory/visual search paradigm that used prelearned categories and highly practiced observers. By contrast, in our Experiment 1 we use CV in an attempt to unravel the mechanisms underlying varied forms of memory search at early stages of practice and with arbitrarily assigned pictures as categories. Although many classic studies of the development of efficient memory and visual search involve paradigms with extensive practice, Nosofsky, Cox, et al. (2014) observed markedly more efficient CM memory search than VM memory search after relatively few practice trials. Thus, obtaining a deeper understanding of the multiple long-term learning mechanisms that may underlie efficient memory search requires examination of performance at early stages of practice as well. Our primary focus in the present set of studies is on these early stage practice effects.

To anticipate, our initial results point strongly in favor of the item-response-learning hypothesis just described, because CV performance involving arbitrary categories is closely aligned with VM performance and is dramatically worse than CM performance. However, our further investigations to test for a role of category learning revealed some factors that complicate the story. Resolving these complicating factors will motivate us to conduct further experiments with more intricate manipulations that will provide a great deal more information concerning the mechanisms that play a role in VM, CM, and CV memory search. We describe these mechanisms in the General Discussion section following our presentation of the complete set of experimental results.

Experiment 1

Each subject was trained in VM, CM, or CV memory search conditions for five blocks of 25 trials each.[1] Set size was 2, 4, or 8, randomly varying from trial to trial.

[1] To explore the role of initial type of training on performance under new item-response-mapping conditions, we then transferred subjects to two blocks of either CM or CV conditions, using the same stimuli as in training. However, the transfer results were difficult to interpret and not the central focus of the present research, so we do not discuss them in this article. In brief, subjects switched to CM versus CV turned out to have, by happenstance, significantly different baseline performances in the original memory search training conditions. The different baselines made it difficult to interpret and draw strong conclusions regarding the patterns of performance in the various transfer conditions.

Method

Subjects. In the first experiment, 175 undergraduate students from Indiana University participated to fulfill an introductory psychology class requirement: 36 in VM, 69 in CM, and 70 in CV. (The different number of subjects in the conditions was due to the transfer manipulation that we mentioned in footnote 1.)

Stimuli. The stimuli were drawn from a pool of 2,400 unique object images obtained from the website of Talia Konkle and described by Brady, Konkle, Alvarez, and Oliva (2008). Each image subtended a visual angle of approximately 7 degrees and was displayed on the center of a gray background. The experiment was conducted with MATLAB's Psychophysics Toolbox (Brainard, 1997) on personal computers.

Procedure. In all conditions, half the tests were targets and half foils. For each subject, 16 stimuli were randomly sampled from the 2,400 images. On each trial in the VM condition a memory set of 2, 4, or 8 items was randomly sampled from the 16-stimulus set. Targets were randomly chosen from the memory set; foils were randomly chosen from the remaining items in the 16-item set. Thus, in the VM condition, items were not grouped into categories and were mapped in varied fashion to target and foil responses. In the CM condition, eight stimuli were randomly drawn from the 16-stimulus set; memory sets were randomly chosen from these same eight on every trial. Targets were randomly chosen from the memory set. Foils were randomly selected from the remaining eight items. Thus, throughout the CM condition, items were grouped in consistent categories and were also mapped consistently to target and foil responses. In the CV condition, two sets of eight stimuli were selected from the 16; these sets were termed *A* and *B*. The memory set on each trial was selected from either *A* or *B*, with *A* or *B* chosen randomly on each trial. A target would be chosen randomly from the memory set, and a foil would be chosen randomly from the eight-item set not providing the memory set. Thus, the items remained grouped in consistent categories throughout the condition but had varied mappings to target and foil responses.

For each trial, a fixation point (asterisk) appeared in the center of the screen for .1 s to indicate the start of that trial. Then each of the memory-set items was presented sequentially for 1 s, followed by .1-s interstimulus intervals. After a 1-s retention interval of a blank screen, a second fixation point (a plus sign) was presented for .5 s, followed by the test probe. The test probe remained on the screen until subjects responded (*J* key = "old," *F* key = "new"). Feedback was then provided on screen for 1 s to indicate whether the response was correct. Subjects were instructed to respond as rapidly as possible while minimizing errors.

Results

The top panel in Figure 1 shows median response times for correct responses for each condition and each set size, averaged across subjects; the lower panel shows probability of an error for each condition and set size, averaged across subjects. Both panels show results averaged across Blocks 2–5 (the first block was considered practice). Data from trials with response times less than 180 ms or greater than 4,000 ms were excluded from these analyses (less than 1% of the data).

Probably due to the use of pictures as stimuli, the error rates in VM were somewhat higher than is usually found when the stimuli are numbers or letters. These error rates are high enough to be considered outcomes of the memory retrieval process (rather than, say, accidental button presses) and should provide evidence consistent with the response time data.

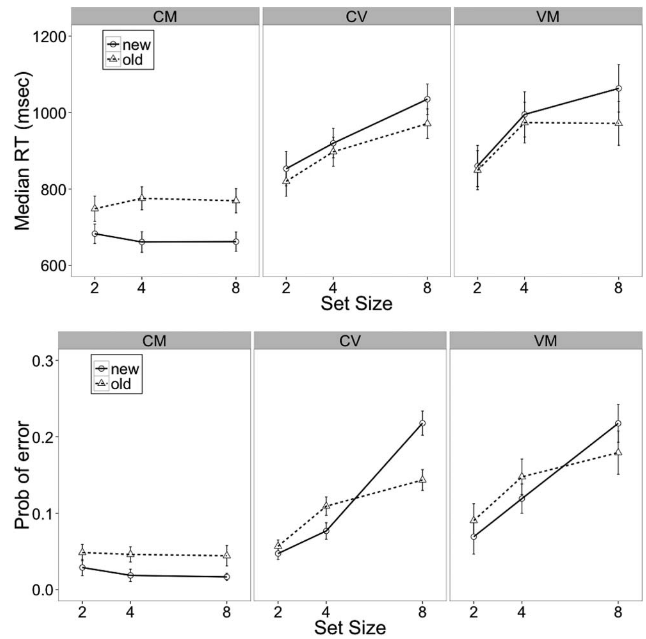

Figure 1. Experiment 1: Median correct response times (RTs; top panel) and probability of error (bottom panel) plotted as a function of condition (CM, CV, and VM), old–new status of probe, and set size. CM = consistent mapping; CV = categorized varied mapping; VM = varied mapping; prob = probability. Bars indicate standard error of the mean.

The data pattern in the CM and VM conditions is consistent with patterns observed in many previous studies, including those in our previous articles (e.g., Nosofsky, Cao, et al., 2014; Shiffrin & Schneider, 1977). CM error rates were low, and CM response times were short. In addition, CM response times varied only slightly (and inconsistently) with changes in set size. VM errors and response times increased strongly with set size. Errors were generally higher and response times much longer than found in CM. Such results were expected and serve as baseline results for comparison with the CV condition.

The middle column in the top and bottom panels of Figure 1 shows the CV data. These data are similar to the VM data in the right column, both showing similar set-size effects; the pattern of results is markedly different from that for CM in the left column. Because the set-size effect is the major measurement of importance for the condition comparisons, we fit linear regression functions to both the response-time data and error-probability data based on set size for each subject. We then performed a 3 × 2 mixed-model analysis of variance (ANOVA) on the set-size function slopes using conditions (CM, CV, and VM) and test-probe types ("old" and "new") as factors. For response time, the condition manipulation was significant, $F(2, 172) = 25.93$, $MSE = 886$, $p < .01$. Pairwise comparisons revealed significant differences in slopes for CM versus CV, $F(1, 138) = 48.99$, $p < .01$, and CM versus VM, $F(1, 102) = 35.85$, $p < .01$. The difference between CV and VM did not approach statistical significance, $F(1, 104) <$

1. For probability of errors, the condition manipulation overall was also significant, $F(2, 173) = 69.36$, $MSE = .00029$, $p < .01$, and the pairwise comparisons showed the same pattern of results as for the RTs: $F(1, 138) = 124.9$, $p < .01$, for CM versus CV, and $F(1, 102) = 93.95$, $p < .01$, for CM versus VM. Again, the CV versus VM difference did not approach statistical significance, $F(1, 104) < 1$.

Discussion

The VM results are standard; they exhibit strong set-size effects, presumably because long-term learning of all sorts is suppressed. Thus, responding is based on storage of the memory set in STM coupled with retrieval that depends on set size (likely because of retrieval processes that depend on lag or serial position; see, e.g., Nosofsky, Cox, et al., 2014). The CM results are also standard (e.g., Schneider & Shiffrin, 1977), albeit perhaps surprising, in that the pattern is found so early in training. It is reasonable to conclude that stimulus-to-response item learning occurred rapidly in CM training, presumably in Block 1 or shortly thereafter, thereby providing a basis for responding independent of set size. Presumably the memory set can be ignored and the response can be based on the learned assignment (foil or target) of each test item.

Most important, because the CV results strongly resemble the VM results, and depart dramatically from the CM results, a reasonable initial conclusion is that categorical encoding has not been used as a basis for responding in CV. It seems likely that categorical encoding has not been learned. (Although it is conceivable that such learning took place but was not used to govern responding, this possibility seems unlikely because subjects usually adopt response methods that minimize cognitive effort.) Finally, the dramatically better performance in the CM condition compared to VM and CV suggests that item-response learning can operate at early stages of training to yield highly efficient CM memory search.

We emphasize that these conclusions are tentative ones, based on our reasoning involving the nature of the CV paradigm and its relation to VM and CM. As will be seen, the next experiment (Experiment 2) reveals some complicating factors involving CV memory search that we had not anticipated. We address these complicating factors in Experiment 3, the results of which confirm the main conclusions reached in Experiment 1 and provide more detailed information concerning the mechanisms that underlie these varied forms of memory search.

Experiment 2A

Whereas Experiment 1 suggests that the greatly enhanced CM performance at early stages of practice is due to item-response learning, in the present experiment we addressed the question: Are categorical codes ever learned and used to govern responding in memory search? One way to assess the use of categorical codes in CV is to employ the use of already well learned categories. Our reasoning is straightforward: If the category codes already exist and do not need to be learned, then subjects should be able to employ the category-coding strategy at the onset of CV training. This approach is the basis for Experiments 2A and 2B, which use letters and numbers. For both CM and CV, the two categories for the stimuli were letters and numbers, whereas the stimuli were all drawn from letters for the VM condition. Our prediction was that CV performance would now strongly resemble CM and depart dramatically from VM.

Logan and Stadler (1991) used the letter and number stimuli in CM but included "catch trials" toward the end of each session: When the memory set was, say, letters, and another letter not in the set occurred as a foil, there was a high tendency to respond "old," consistent with a categorical basis for responding. Our Experiment 2 may provide converging evidence for Logan and Stadler's conclusions regarding the use of category coding.

Method

Subjects. In this experiment, 70 undergraduate students from Indiana University participated to complete course credit. There were 23 subjects in the CM condition, 24 subjects in the CV condition, and 23 subjects in the VM condition.

Stimuli. The stimuli were the English alphabet set and single-digit number set, excluding *O, I, 1,* and *0* to avoid confusion. The letters were all capital letters. The stimuli were presented in the center of the computer screen in Courier font size 40.

Procedure. In all conditions, subjects each had seven blocks of training, with 25 trials per block. The procedures matched those in Experiment 1 in most respects—set sizes were again 2, 4, and 8. In the CM and CV conditions, a letter-category set was created by randomly selecting eight letters from the total letter set for each subject. The number-category set was the same for every subject and contained the numbers two to nine. In the VM condition, a set of 16 letters was randomly selected from the letter set for each subject.

In the VM condition, on each trial, the memory set was randomly sampled from the 16-item set and the foil randomly sampled from the remaining letters. In the CM condition, the memory set was always randomly sampled from the letter-category set. Foils were randomly sampled from the number-category set. In the CV condition, on each individual trial, the memory set was taken from either the letter category or the number category. If the test item was a foil, then it was chosen from the other set. All other aspects of the procedure were the same as in Experiment 1.

Results

Although we expected memory search for letters and/or numbers to be easy, that did not turn out to be case for all subjects. We explored various criteria for eliminating poorly performing subjects—none changed the pattern of results, but the least noisy results were found when we used a criterion that eliminated the worst performing six subjects in CV, three subjects in VM, and two subjects in CM—overall accuracy less than 0.7 and/or median response time (RT) greater than 1,200 ms in the CV and VM conditions; median RT greater than 1,000 ms for the CM condition. For the remaining subjects, there were no trials with response times shorter than 180 ms or longer than 4,000 ms. As in Experiment 1, we considered Block 1 as practice and next report the results for all the subsequent blocks.

The results are shown in Figure 2 using the same format as for Experiment 1. CM and VM showed the usual patterns, with the set-size functions being nearly flat for CM but increasing steeply for VM. To our surprise, however, rather than strongly resembling

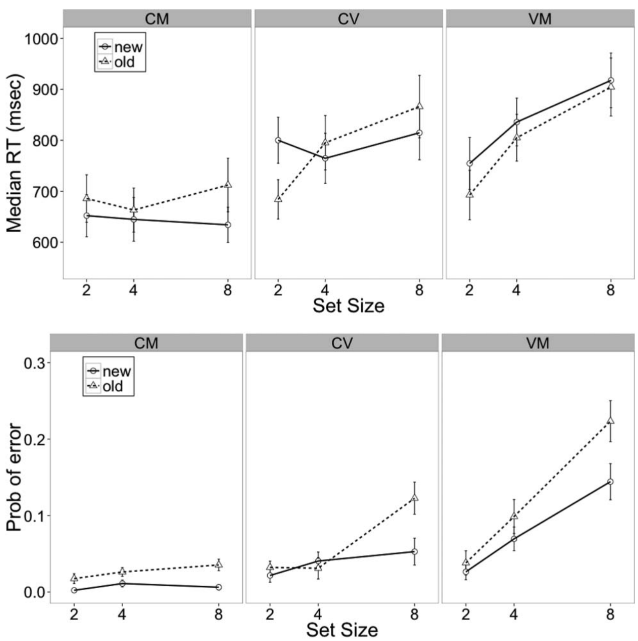

Figure 2. Experiment 2A: Median correct response times (RTs; top panel) and probability of error (bottom panel) plotted as a function of condition (CM, CV, and VM), old–new status of probe, and set size. CM = consistent mapping; CV = categorized varied mapping; VM = varied mapping; prob = probability. Bars indicate standard error of the mean.

CM, a better summary is that the CV results were overall intermediate between CM and VM performance. Indeed, for the RTs, the CV results seem more similar to those for VM than for CM (although the RT data were noisy). The error data seem more similar overall to those for CM than for VM, but even here there was a clear set-size effect for the old items. We again performed a 3 × 2 mixed-model ANOVA on the set-size function slopes using conditions (CM, CV and VM) and test-probe types (old and new) as factors. Overall, the condition manipulation was significant for both RTs and error rates: RT: $F(2, 56) = 12.13$, $MSE = 641$, $p < .01$; probability of error: $F(2, 56) = 33.73$, $MSE = .00017$, $p < .01$. Pairwise comparisons indicated that CV was significantly different from both CM and VM for probability of errors: CV vs. CM: $F(1, 37) = 15.25$, $p < .01$; CV vs. VM: $F(1, 36) = 15.94$, $p < .01$. For response time, CV is significantly different from CM, $F(1, 37) = 8.94$, $p < .01$, but not from VM, $F(1, 36) = 2.14$, $p = .15$. As expected, performance for CM was significantly different from that for VM for both response time, $F(1, 39) = 71.04$, $p < .01$, and probability of errors, $F(1, 39) = 36.59$, $p < .01$.

These results indicate some difficulty in using the well-learned categories of letters and numbers to govern responding in CV, at least for some subjects and some trials. It is possible that the switching of response assignments from trial to trial is confusing and inhibits efficient use of a categorical strategy for responding. An analysis by blocks was quite noisy but seemed to indicate a trend for the CV results to move farther from the VM pattern as blocks continued.

We decided, therefore, to conduct a similar study with longer periods of CV training. Four full sessions of CV training were followed by two sessions of CM training with the same stimuli.

Experiment 2B

We tested four subjects using the same procedure as in the CV conditions of Experiment 2A. The subjects were tested for four sessions of seven blocks of 25 trials. The subjects then switched to two sessions of CM, always drawing memory sets from each subject's letter category set. Subjects were paid $12 for the completion of each session.

Results

Figure 3 shows median response times (averaged across subjects) for correct responses as a function of sessions, test-probe type (old vs. new), and set size—the first four panels CV and the last two CM. Because error rates were low (less than .05) for all set sizes in all conditions, we do not show plots of the error functions. The Session 1 RT results were much like the intermediate CV results in Experiment 2A. However, Sessions 2, 3, and 4 show results much like those for CM. In fact, the switch to CM for Sessions 5 and 6 did not produce dramatic changes in performance (RTs got slightly shorter, but error rates slightly increased).

Discussion

Why there should have been such a marked change in CV performance, especially in the first two sessions, is not clear. Perhaps it takes time to become accustomed to the constant switching from trial to trial. Perhaps it takes time to realize that an alternative and more efficient response strategy is available. Whatever the reasons for the slow switch, it is clear that categorical coding as a strategy for performing CV memory search task occurred with sufficient training (at least in cases involving prelearned categories).

Nevertheless, the Session 1 results involving CV memory search with these prelearned categories leave somewhat unclear the interpretation of our original Experiment 1 results. There, we found that at early stages of practice, CV performance involving arbitrary category sets closely resembled VM performance and departed dramatically from CM performance. Our interpretation was that

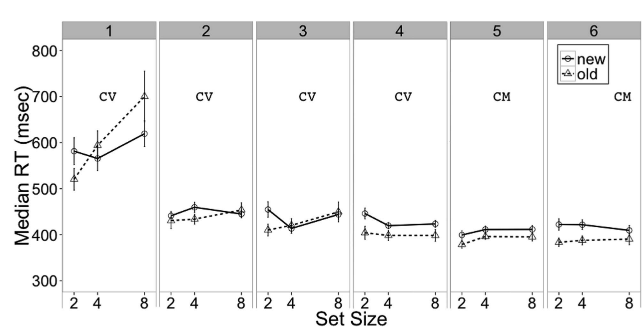

Figure 3. Experiment 2B: Median correct response times plotted as a function of sessions (CV: 1–4, CM: 5–6), condition (CV, CM), old–new status of probe, and set size. CM = consistent mapping; CV = categorized varied mapping. Bars indicate standard error of the mean.

subjects had not learned codes for the arbitrary category sets and that item-response learning was the fundamental basis for the enhanced CM performance. However, in Experiment 2, we found that even with prelearned categories for which category codes were readily available, CV performance departed significantly from CM performance at the early stages of practice. Thus, a logical possibility is that subjects were able to rapidly learn category codes even for arbitrary sets in both CV and CM search and that the early CV deficits arose because of the continual switching of response assignments.

To disentangle these multiple factors and gain a deeper understanding of their joint role in memory-search performance, we conducted Experiment 3, which attempted to control for switching costs while also manipulating memory-search conditions (VM, CM, CV) and whether the categories are arbitrary or prelearned.

Experiment 3

One aim of Experiment 3 was to continue to use the CV paradigm as a vehicle for contrasting item-response learning versus category learning in memory search, while controlling for the switching costs that seem to arise during the early stages of CV practice. A second aim was to compare memory search involving arbitrary and prelearned categories within the same experiment and using the same types of stimuli. As is seen in this section, these manipulations allowed us to draw stronger conclusions regarding the long-term learning mechanisms that contribute to performance in the memory-search task.

To control for switching costs, we conducted a "blocked" version of the CV task in which the target set and foil set were held fixed for 10 consecutive trials. The assignment of sets to target and foil responses was then switched and held fixed for the next 10 consecutive trials, and so on, back and forth. Thus, just as in our previous experiments, there were still two fixed categories of items throughout the entire experiment, allowing for category coding in the CV condition. In addition, just as in our previous experiments, long-term item-response learning could take place because the members of each category were assigned equally often to target and foil responses. However, unlike in our previous experiments, within each 10-trial sequence, there was ample opportunity for subjects to become accustomed to which category was aligned with target responses and which category was aligned with foil responses (if such category coding had indeed occurred). We refer to this condition as *CV-blocked.*

For purposes of comparison, we again tested a CM condition as well as a VM condition. However, to gauge the extent to which short-term forms of item learning can contribute to performance, we tested two versions of the VM condition. One version was the standard version in which the target and foil sets were randomly selected anew on each individual trial. The second version was a *VM-blocked* condition. Here, the target and foil sets remained constant across 10-trial blocks, then were randomly resampled for each subsequent 10-trial block (see the Method section for details). The upshot is that there were no long-term categories that could be learned in the VM-blocked condition, but short-term memory for recently presented items was equated across the CV-blocked and VM-blocked conditions.

Finally, in both the CV-blocked and CM tasks, we tested conditions involving arbitrary category sets as well as prelearned category sets. Holding stimulus types fixed across conditions, the prelearned categories were pictures of small animals versus small human-made objects. Naturally, we expected that subjects could make easy use of category codes in the conditions involving prelearned categories. Thus, assuming that our blocking manipulation would remove switching costs, then the prediction was that the set-size functions should be nearly flat in the CV-blocked condition with prelearned categories, even during the early stages of practice. Furthermore, CV-blocked performance with prelearned categories should resemble CM performance. The critical question concerned the nature of the -size functions in the CV-blocked condition with the arbitrary categories. If category coding developed early, then the set-size functions should be flat in the CV-blocked condition with arbitrary categories as well. However, if item-response learning is the main basis for enhanced CM performance with arbitrary categories (at early stages of practice), then performance in the CV-blocked condition with arbitrary categories should instead resemble VM-blocked performance. Finally, we could gauge the extent to which STM for recently presented items from previous lists might contribute to performance by comparing CV-blocked and VM-blocked to standard VM.

Method

Subjects. The subjects were 114 undergraduate students from Indiana University who participated in partial fulfillment of an introductory psychology course requirement.

Stimuli. The stimuli consisted of 32 unique pictures of small animals and small man-made objects (16 of each category) from Konkle's website and described in Konkle and Caramazza (2013). (The website contains 60 pictures from each category. We chose 16 of each with the aim of including distinct subtypes of each category as well as eliminating certain animals that seemed likely to arouse emotional reactions, such as snake, rat, and spider). Each image subtended a visual angle of approximately 7 degrees and was displayed in the center of a gray background. The experiment was conducted with MATLAB Psychophysics Toolbox (Brainard, 1997) on personal computers.

Procedure. Subjects were randomly assigned to one of the six conditions described in the introduction to this experiment (CV-blocked arbitrary, CV-blocked prelearned, CM arbitrary, CM prelearned, VM-blocked, and VM standard). In all conditions, (a) there were five blocks of trials with 30 trials per block; (b) memory-set size was 2, 4, or 8, randomly determined on each trial; and (c) the type of test probe (old vs. new) was also randomly determined on each trial.

In the conditions involving the prelearned categories (CM prelearned and CV-blocked prelearned), for each subject, eight items from the object set and eight items from the animal set were randomly selected. In the CM condition, for each subject, either the object set or the animal set was randomly chosen to serve as the target set throughout the experiment. The other set would serve as the foil set. The CM procedure was then just as described in our previous experiments.

At the start of the CV-blocked (prelearned) condition, either the object set or the animal set would be randomly chosen to serve as the target set, with the other set serving as the foil set. This assignment stayed fixed for 10 trials. The assignment of categories to target and foil sets would then switch and be held constant for the next 10 trials, and so on for the remainder of the experiment. In all other respects, the CV-blocked condition was the same as described in our previous experiments.

In the conditions involving the arbitrary categories (CM arbitrary and CV-blocked arbitrary), the stimuli were all chosen from one of the prelearned categories (objects or animals), randomly determined for each subject. Half the items were randomly assigned to Category A and the other half to Category B. In the CM condition, Category A would serve as the target set and Category B as the foil set throughout the experiment. In the CV-blocked condition, the assignment of the A and B items to the target and foil sets would switch every 10 trials, just as described earlier for the CV-blocked prelearned condition. We decided to use a single prelearned category to form each of the A and B sets (rather than randomly sampling from both categories) so that subjects would not have strong prior beliefs about which items should be grouped together.[2]

In the VM conditions, one of the prelearned categories was randomly chosen to serve as the full set of stimuli throughout the experiment. At the start of the VM-blocked condition, eight of the 16 items were randomly assigned as Set A, and the remaining eight items as Set B. Set A would serve as the target set, and Set B as the foil set for 10 consecutive trials. In the next block of 10 trials, a new set of randomly sampled items would serve as Set A and the remaining items as Set B. (Thus, there were no permanent arbitrary categories; instead, although the target and foil sets remained fixed for 10 trials, their composition changed randomly in each subsequent sequence of 10 trials.) To remove accidental runs of near-permanent categories, we constrained the sampling such that half the items assigned as targets in one block would serve as foils in the next block. Finally, in the VM-standard condition, the memory set was randomly sampled from the full set on each trial, and a foil would be randomly sampled from all remaining items. A schematic illustration of the full set of conditions is provided in Figure 4.

Other details of the procedure were the same as described in Experiment 1. Subjects received no explicit instructions about the 10-trial sequences in the CV-blocked and VM-blocked conditions.

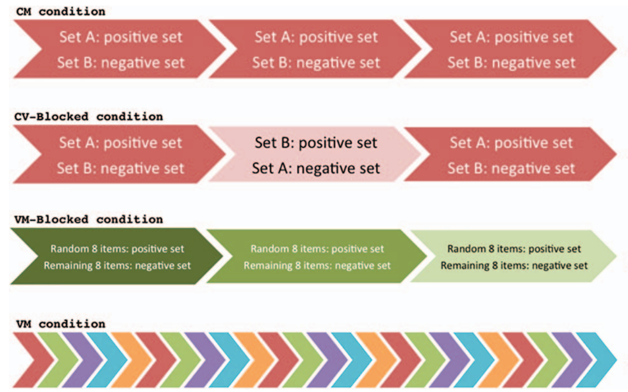

Figure 4. A schematic illustration of the conditions in Experiment 3. In both the blocked CV and blocked VM conditions, each color block represents a 10-trial sequence where memory sets are drawn from a positive set that is fixed through the 10 trials and the set then changes for the next 10 trials. In the CM condition, the positive set was fixed across all trials. In the VM condition, the set randomly changed for every trial. CM = consistent mapping; CV = categorized varied mapping; VM = varied mapping. See the online article for the color version of this figure.

Results

The first block was considered practice and was excluded from all the following data analyses. Data from trials with RTs less than 180 ms or greater than 4,000 ms were also excluded from these analyses (1.2% of the data).

To begin, we first examined performance in the CV-blocked condition with prelearned categories as a function of trials in the 10-trial sequences. These results are plotted in Figure 5, separately for the new and old test probes, with the top panel showing the median RTs and the bottom panel the error probabilities, averaged across subjects. Although performance for the old test probes did not change systematically across the 10-trial sequences, performance on the new test probes improved, particularly from the first three trials to the following trials. Based on these results, we decided to exclude the first three trials of each 10-trial sequence in all remaining analyses across all conditions.

The median correct RTs are plotted as a function of condition, test-probe type (old vs. new), and set size in Figure 6. The error data are plotted in Figure 7. In both figures, the results from the conditions involving arbitrary categories are plotted in the left columns, whereas the results involving the prelearned categories are plotted in the right columns.

To begin, note that performance in the CM conditions was clearly better than in all remaining conditions and that the CM set-size functions are generally nearly flat. (In cases in which there are slight changes in CM performance with set size, the changes are inconsistent across the new and old items.) Second, as expected, overall performance was worst in the standard VM condition, with the set-size functions for both RTs and errors steeply increasing. The set-size functions also increased in the VM-blocked condition, although performance in this condition was somewhat better than in the standard VM condition (see the discussion later).

The key results of interest involve performance in the CV-blocked conditions and how performance in these conditions related to performance in the other conditions. First, note that in the case involving prelearned categories (right panels), the set-size functions in the CV-blocked condition are flat for the new items and increase only slightly for the old items. Apparently, once switch costs were reduced through use of the blocking procedure, subjects were able to make far better use of the category-coding strategy as a basis for performing the task (compared to the early stages of practice in Experiment 2). By contrast, in the case involving the arbitrary categories, the set-size functions in the CV-blocked condition are generally steeply increasing (the single exception is the error-probability function for new items). Indeed, in the case involving arbitrary categories, performance

[2] A technical concern is the level of features that serve to discriminate the prelearned categories and the arbitrary categories. A category code can be regarded as a "high level" feature. It is conceivable that low-level perceptual features might be available that separate the members of the prelearned categories. No salient low-level features were apparent to us upon visual inspection of the stimuli. Future research would be needed to investigate this issue.

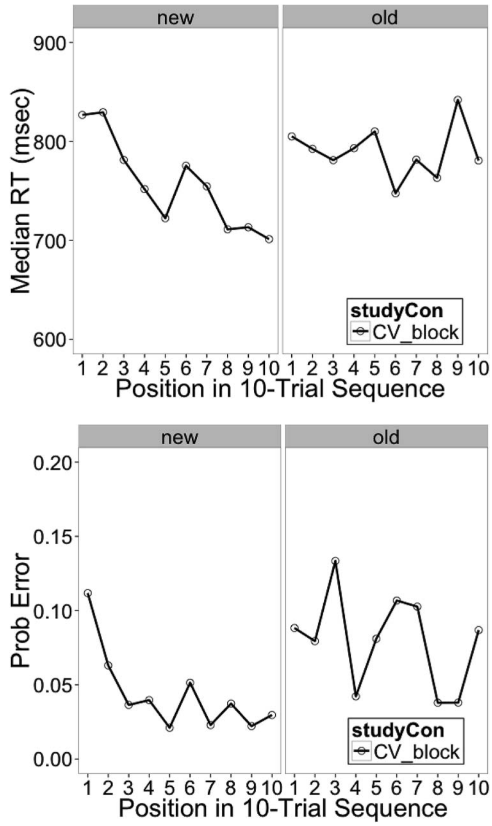

Figure 5. Experiment 3: Median correct response times (RTs; top panel) and probability of errors (bottom panel) plotted as a function of old–new status of probe and position in the 10-trial sequences in the CV-blocked condition with prelearned categories. CV = categorized varied mapping; studyCon = study condition; prob = probability.

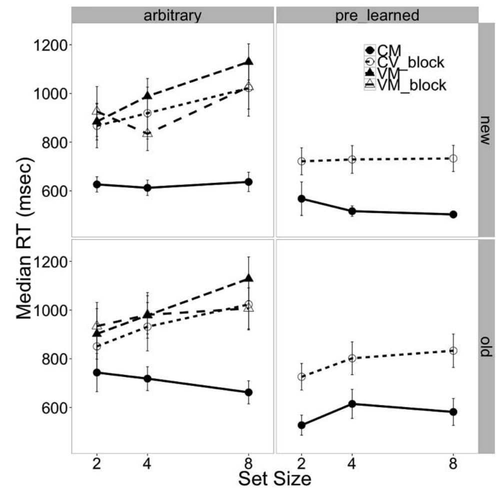

Figure 6. Experiment 3: Median correct response times (RTs) plotted as a function of condition (CM, blocked CV, VM, blocked VM), category type (arbitrary vs. prelearned), old–new status of probe, and set size. Bars indicate standard error of the mean across subjects. CM = consistent mapping; CV = categorized varied mapping; VM = varied mapping.

in the CV-blocked condition departed dramatically from CM performance and was quite similar to performance in the VM-blocked condition. Thus, it appears that observers were unable to use a category-coding strategy at these early stages of practice in the CV-blocked condition when the sets involved arbitrary categories. A reasonable inference, therefore, is that the greatly enhanced early practice performance in the CM condition in cases involving arbitrary categories is based on item-response learning.

We conducted statistical analyses to confirm our descriptions of the comparisons of performance across the arbitrary and prelearned conditions of the CV-blocked task. In these analyses, we fitted least-squares regression lines to the set-size functions yielded by each individual observer in each condition, separately for the old and new test probes. The mean regression-line slopes are displayed in Figure 8: The left panel displays the results for the RTs, and the right panel displays the results for the error probabilities. Except for the case involving error probabilities for new items, the mean set-size slopes are greater for the arbitrary categories than for the prelearned ones, and the prelearned slopes are generally near zero. We conducted 2 × 2 ANOVAs on the individual-subject set-size slope data, using condition (arbitrary vs. prelearned) and test-probe type (old vs. new) as factors. There was a main effect of condition for the RT data, $F(1, 36) = 8.63$, $MSE =$ 695, $p < .01$, reflecting the steeper slopes in the arbitrary condition compared to the prelearned condition. There was no main effect of condition for the error-probability data, $F(1, 36) = 2.418$, $p = .129$, reflecting the lack of any effect for the new items. A focused t test for the old items was significant in a one-tailed test: $t(36) = 1.71$, $p = .048$.

As we noted earlier, performance in the VM-blocked condition was better than performance in the standard VM condition (and performance in the CV-blocked condition with arbitrary categories was similar to that for the VM-blocked performance). Because

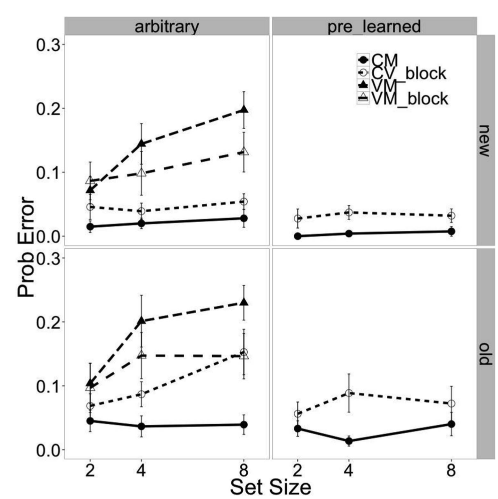

Figure 7. Experiment 3: Probability of error plotted as a function of condition (CM, blocked CV, VM, blocked VM), category type (arbitrary vs. prelearned), old–new status of probe, and set size. Bars indicate standard error of the mean across subjects. CM = consistent mapping; CV = categorized varied mapping; VM = varied mapping; prob = probability.

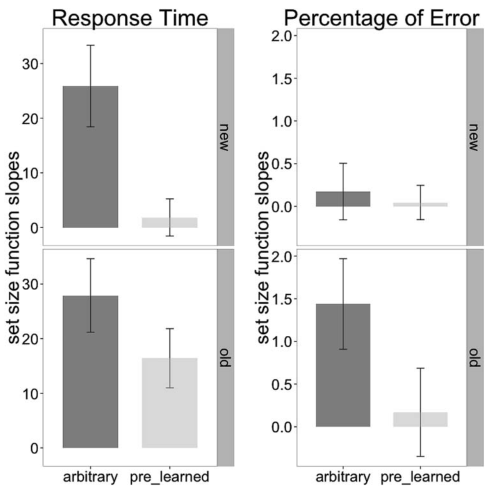

Figure 8. Comparison of subjects' performance for arbitrarily assigned categories (dark bars) and prelearned categories (light bars). Left panels: Average set-size slopes of correct response times (msec) in the CV-blocked conditions. Right panels: Average set-size slopes of percentage of error in the CV-blocked conditions. Error bars indicate standard error of the mean across subjects. CV = categorized varied mapping.

there were no long-term categories to be learned in the VM-blocked condition and there was no possibility of long-term item-response learning, it is apparent that the better performance in this condition was due to memory for recent lists and test probes. In the VM-blocked condition, within each 10-trial sequence, items maintained consistency in their assignments as positive versus negative probes. Thus, there seem to be two related hypotheses to explain the advantages of blocking in VM: (1) Item learning occurs rapidly and is used in the latter stages of each block and (2) in nonblocked VM, interference arises from immediately preceding trials that have differing response assignments for the current-trial test item; blocking removes this interference. The present evidence does not provide a clear way to distinguish these hypotheses—although item learning can occur quite quickly, it is not clear whether such learning continues when the response assignments switch every 10 trials.

A final point to note is that, even in the case involving prelearned categories, overall performance in the CM condition was better than in the CV-blocked condition (see the right panels of Figures 6 and 7), although the set-size functions in both conditions were nearly flat. It seems to be the case that consistent mappings without switches provide an overall response-time benefit. This result is consistent with a finding by Shiffrin and Schneider (1977, Experiment 3); they found that for highly practiced observers (roughly 24 sessions of training) engaging in hybrid visual/memory search, CM performance was better than CV performance.

General Discussion

Summary

The present research aimed to explore the combined, interactive roles of short-term and long-term memory in probe-recognition memory search. This issue was studied by exploring the long-term learning mechanisms that lead to the highly efficient memory search (flat set-size functions) observed under conditions of consistent-mapping (CM) training. The basic idea is that the normal processes of checking the just-studied list of items held in STM is supplemented by response mappings that are learned over time and held in long-term memory. We distinguished between two potential forms of long-term learning: item-response learning versus category learning. In item-response learning, subjects learn long-term mappings between individual items and target versus foil responses. In category learning, subjects learn high-level codes corresponding to separate sets of items and learn to attach old versus new responses to these category codes. In an attempt to distinguish between these two forms of learning, we tested subjects in categorized varied mapping (CV) conditions: There were two distinct categories of items, but the assignment of categories to target versus foil responses varied across trials. Our reasoning was that if the basis for efficient performance in the CM condition relies solely on category learning, then CV performance should closely resemble CM performance. By comparison, if item-response learning is playing the major role, then CV performance should resemble performance in standard varied-mapping (VM) conditions, with CM performance being dramatically better than both. Our subsequent experiments, however, revealed that some more nuanced mechanisms are also involved.

We tested between these alternatives in three separate experiments that used varieties of different stimuli, types of categories, amount of training, and sequential relations among study-test lists. In Experiment 1, the categories were composed of novel, arbitrarily assigned pictures, and we investigated performance during a single session of testing. CV performance departed dramatically from CM performance and strongly resembled VM performance, providing initial support for the item-response-learning hypothesis. In Experiment 2, we used prelearned categories composed of letters versus digits for which it should be easy to use a category-coding strategy to perform the CV memory-search task. Unexpectedly, however, we found that at early stages of practice, CV performance remained worse than CM performance (although better overall than VM performance). Apparently, the costs of switching the assignments of categories to target and foil responses from trial to trial interfered with the use of the category-coding strategy. Because the same costs would be expected to have occurred in Experiment 1, the original support for the item-response-learning hypothesis in that earlier experiment was rendered somewhat ambiguous.

Therefore, in Experiment 3, we conducted manipulations designed to greatly lessen the CV switching costs by blocking across 10-trial sequences the assignment of categories to target and foil responses. In addition, in this experiment, we held constant the stimulus types across conditions and simultaneously manipulated whether the memory-search task involved arbitrary versus prelearned categories. In the condition involving prelearned categories, CV performance now yielded flat set-size functions that resembled those for CM performance. By contrast, in the condition involving arbitrary categories, CV performance again departed dramatically from CM performance and yielded steep set-size functions that resembled those for VM performance. Thus, in this latter condition, it appears that at the early stages of practice, subjects did not form and use category codes corresponding to the two separate, arbitrary sets of objects. A reasonable inference, therefore, is that at early stages of practice involving CM memory search with arbitrary sets, the greatly enhanced performance is indeed due to item-response learning and not category learning.

Relations to Other Work

Our results suggest that item-response learning plays a fundamental role in CM memory search and that it occurs rapidly. For example, in our Experiment 3, just a few trials with consistent item mappings led to clear benefits in both VM and CV performance with arbitrary categories. A similar form of rapid item learning appears to underlie repetition-priming effects in lexical decision (Logan, 1990) and in a variety of classification tasks such as relative-size classification (Dobbins, Schnyer, Verfaellie, & Schacter, 2004; Horner & Henson, 2011). In a sample version of the relative-size classification paradigm, subjects were instructed to press one of two keys to indicate whether a test object was larger than a reference object. Following initial testing, there was a transfer phase in which the reference object was changed such that the response judgment remained the same for some items (congruent) but switched for other items (incongruent). The key finding was that response times for congruent stimuli are much faster compared to novel stimuli, whereas the effect is much smaller or even reverses for incongruent stimuli (Horner & Henson, 2011). A general interpretation is that these repetition-priming effects arise from forms of rapid item-response learning as opposed to a general facilitation of visual classification processes (because general facilitation would lead to better performances for both congruent and incongruent stimuli). The results from these repetition-priming classification tasks are consistent with our results, and it is possible that a similar item-response learning mechanism accounts for the rapid performance improvement in both tasks.

Subjects were slow to adopt the category-coding strategy in the CV condition for well-learned categories in our Experiment 2. Whereas item-response learning allows observers to form reliable responses for individual items, the category-coding strategy could lead to overgeneralization and produce false-alarm responses to other stimuli in the category. Therefore, observers might be reluctant to adopt the category-coding strategy at early learning stages and utilize the strategy only when confident. For example, in Experiment 3 of Logan and Stadler (1991), subjects received CM training for well-learned categories (letters vs. numbers) in a memory-search paradigm. At different stages of training, a catch trial was inserted: A novel member of the target-set category was presented as a new test probe (e.g., if targets were always letters, a novel letter would be presented as a foil). Performance on standard trials rapidly improved, but the false-alarm rate to catch trials increased more slowly as the training progressed. This result is consistent with the hypothesis that subjects learn and use item responses quickly but learn and use category coding more slowly. It is of course plausible that there are switch costs associated with the training in our Experiment 2 (as suggested by the results from our Experiment 3) and that could also have slowed the adoption of a category-response strategy.

Future Research Directions

With sufficient practice or with designs that remove switching costs, it is apparent that subjects can use the category-coding strategy in CV tasks involving prelearned categories (Experiments 2 and 3). We hypothesize that the same pattern would eventually be observed for the arbitrarily assigned picture stimuli once sufficient category learning took place. It may be that category learning for arbitrary groups of stimuli occurs more slowly than does item-response learning because category learning is more abstract and indirect, because the arbitrarily assigned stimuli belong to other preexisting natural categories or, more likely, both.

We should acknowledge that our logic involving the use of CV for contrasting item-response versus category learning is not foolproof. Our assumption has been that in both CM and CV, the two sets of items assigned to target and foil response are always separated, so any learning that groups the items into categories should proceed at roughly the same pace across the conditions. It is possible to argue, however, that, for some reason, the asymmetric method of separating the categories in CM, with one category of items experienced frequently in the memory sets and the second set experienced infrequently as foils, leads to more efficient category learning than occurs in CV. Future research would be needed to investigate this possibility.

Another goal of future research is to obtain converging evidence for our conclusions regarding the role of long-term learning in CM memory search. One approach that we find particularly intriguing was reported by Carlisle, Arita, Pardo, and Woodman (2011). These researchers conducted a paradigm in which observers engaged in visual search for a single target that remained fixed across trials. The expectation was that there would be a shift from a visual working memory representation to a long-term memory representation for this single consistently mapped target. Carlisle et al. (2011) observed a neurological signature of this shift by measuring changes in the contralateral-delay activity of event-related potentials across trials of the visual-search task. Perhaps such an approach could provide evidence of the forms of long-term item-response learning that we hypothesize occurs in the context of the present memory-search tasks that involve much larger sets of consistently mapped targets.

Finally, it is important that formal models be developed that capture the joint influence of item-response learning and category coding in the development of automaticity in memory search. There are a number of memory-search models designed to account for how the history of experience with previous lists influences performance on current lists (e.g., Banks & Atkinson, 1974; Logan, 1988, 1990; Nosofsky, 2016; Nosofsky, Cao, et al., 2014; Nosofsky, Cox, et al., 2014). The emphasis in such models is on evidence-accumulation processes driven by item retrieval. However, in the case of prelearned categories, and perhaps in the case of arbitrary categories with highly practiced subjects, such models would need to be extended to account for the role of category coding in CV and CM memory search. It is an open question whether the category-coding process might be modeled as simply contributing to the same evidence-accumulation process that is driven by retrieved items or whether some mixture-of-strategies model might be needed to account for the data.

Conclusion

In sum, it appears that two mechanisms of long-term learning may contribute to the efficient performance observed in consistent-mapping memory search—item learning and category learning—

and that the contributions of each mechanism follow different time courses and vary with experimental conditions. Item learning is the main driver of performance at early stages of practice involving arbitrary categories, whereas category-coding strategies play an important role in cases involving prelearned categories. It is an open question whether category-coding processes may also contribute in cases involving arbitrary sets at later stages of practice. These results add to the theoretical understanding of the mechanistic bases for CM performance, one of the hallmark examples of the development of forms of automaticity in cognition.

References

Banks, W. P., & Atkinson, R. C. (1974). Accuracy and speed strategies in scanning active memory. *Memory & Cognition, 2,* 629–636. http://dx.doi.org/10.3758/BF03198131

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105,* 14325–14329. http://dx.doi.org/10.1073/pnas.0803390105

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. http://dx.doi.org/10.1163/156856897X00357

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience, 31,* 9315–9322. http://dx.doi.org/10.1523/JNEUROSCI.1097-11.2011

Cheng, P. W. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review, 92,* 414–423. http://dx.doi.org/10.1037/0033-295X.92.3.414

Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004, March 18). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature, 428,* 316–319. http://dx.doi.org/10.1038/nature02400

Horner, A. J., & Henson, R. N. (2011). Stimulus-response bindings code both abstract and specific representations of stimuli: Evidence from a classification priming design that reverses multiple levels of response representation. *Memory & Cognition, 39,* 1457–1471. http://dx.doi.org/10.3758/s13421-011-0118-8

Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience, 33,* 10235–10242. http://dx.doi.org/10.1523/JNEUROSCI.0983-13.2013

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527. http://dx.doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22,* 1–35. http://dx.doi.org/10.1016/0010-0285(90)90002-L

Logan, G. D., & Stadler, M. A. (1991). Mechanisms of performance improvement in consistent mapping memory search: Automaticity or strategy shift? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 478–496. http://dx.doi.org/10.1037/0278-7393.17.3.478

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General, 118,* 346–373. http://dx.doi.org/10.1037/0096-3445.118.4.346

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10,* 465–501. http://dx.doi.org/10.1016/0010-0285(78)90008-7

Nosofsky, R. M. (2016). An exemplar-retrieval model of short-term memory search: Linking categorization and probe recognition. *Psychology of Learning and Motivation, 65,* 47–84. http://dx.doi.org/10.1016/bs.plm.2016.03.002

Nosofsky, R. M., Cao, R., Cox, G. E., & Shiffrin, R. M. (2014). Familiarity and categorization processes in memory search. *Cognitive Psychology, 75,* 97–129. http://dx.doi.org/10.1016/j.cogpsych.2014.08.003

Nosofsky, R. M., Cox, G. E., Cao, R., & Shiffrin, R. M. (2014). An exemplar-familiarity model predicts short-term and long-term probe recognition across diverse forms of memory search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1524–1539. http://dx.doi.org/10.1037/xlm0000015

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review, 118,* 280–315. http://dx.doi.org/10.1037/a0022494

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66. http://dx.doi.org/10.1037/0033-295X.84.1.1

Schneider, W., & Shiffrin, R. M. (1985). Categorization (restructuring) and automatization: Two separable factors. *Psychological Review, 92,* 424–428. http://dx.doi.org/10.1037/0033-295X.92.3.424

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190. http://dx.doi.org/10.1037/0033-295X.84.2.127

Sternberg, S. (1966, August 5). High-speed scanning in human memory. *Science, 153,* 652–654. http://dx.doi.org/10.1126/science.153.3736.652

Sternberg, S. (2016). In defense of high-speed memory scanning. *Quarterly Journal of Experimental Psychology, 69,* 2020–2075. http://dx.doi.org/10.1080/17470218.2016.1198820

Wolfe, J. M., Boettcher, S. E., Josephs, E. L., Cunningham, C. A., & Drew, T. (2015). You look familiar, but I don't care: Lure rejection in hybrid visual and memory search is not based on familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 1576–1587. http://dx.doi.org/10.1037/xhp0000096

Received May 27, 2016
Revision received September 7, 2016
Accepted September 29, 2016 ■